**Claims:**

1. A method of producing, in yeast, a hydroxylated triple helical protein, said method comprising the steps of:

(A) introducing into a suitable yeast host cell:

(i) a first DNA molecule comprising a DNA sequence encoding prolyl 4-hydroxylase I-subunit (P4HI) operably linked to a promoter functional in said yeast host cell,

(ii) a second DNA molecule comprising a DNA sequence encoding prolyl 4-hydroxylase 9-subunit (P4H9) operably linked to a promoter functional in said yeast host cell, and

(iii) a third DNA molecule comprising a DNA sequence encoding a polypeptide or peptide operably linked to a promoter functional in said yeast host cell, wherein said polypeptide or peptide is one which, when hydroxylated, forms said hydroxylated triple helical protein, and wherein said polypeptide or peptide is a synthetic polypeptide or peptide represented by the following formula:

$$(A)_l - (B)_m - [Z] - (C)_o - (D)_p,$$

wherein;

Z is a domain comprising two or more repeat units of the formula:

$$[(E)_q - (GlyXY)_i - (F)_r],$$

wherein;

E and F represent sequences of one or more amino acids, which sequences may vary from repeat unit to repeat unit, and for each repeat unit q and r are each independently selected from 0 and 1, and

i is $\geq 1$ such that domain Z comprises 2 to 1500 GlyXY triplets,

Gly represents glycine, and

X and Y, which may be the same or different, represent an amino acid, and wherein the identity of each amino acid represented by X and Y may vary from GlyXY triplet to GlyXY triplet, but wherein at least one Y of the $(GlyXY)_i$ sequence must be proline,

A and D, which may be the same or different, each represent a polypeptide or peptide domain which optionally comprises a triple helical forming repeating sequence $(GlyXY)_n$, and l and p are each independently selected from 0 and 1,

B and C, which may be the same or different, each represent a polypeptide or peptide domain which is heterologous to collagen proteins and which does not comprise a triple helical forming repeating sequence $(GlyXY)_n$, and m and o are each independently selected from 0 and 1; and

(B) culturing the resulting yeast host cell of step (A) under conditions suitable to express said P4HI and P4Hϑ and said synthetic polypeptide or peptide, to produce said hydroxylated triple helical protein;

wherein during culturing in step (B), each of said first DNA molecule, said second DNA molecule and said third DNA molecule are replicated, stably retained and segregated by the yeast host cell.

2. The method of claim 1, wherein domain Z comprises no more than 10 to 300 GlyXY triplets.

3. The method of claim 1, wherein in domain Z, $(GlyXY)_i$ has an amino acid length which is at least three times greater than the combined amino acid length of E and F.

4. The method of claim 1, wherein expression of the said P4HI subunit and said and P4Hϑ subunit is controlled in a coordinated manner by a bidirectional promoter.

5. The method of claim 4, wherein said bidirectional promoter is yeast GAL1-10 promoter sequence.

6. The method of claim 1, wherein said P4HI subunit is an avian P4HI subunit or a mammalian P4HI subunit, and said P4Hϑ subunit is an avian P4Hϑ subunit or a mammalian P4Hϑ subunit.

7. The method of claim 6, wherein said mammalian P4HI subunit is human P4HI subunit, and said mammalian P4Hϑ subunit is human P4Hϑ subunit.

8. The method of claim 1, wherein each of said second DNA molecule and said third DNA molecule further comprise a DNA sequence encoding a secretion signal

such that said P4H9 and said polypeptide(s) or peptide(s) are expressed and secreted by said yeast host cell.

9. The method of claim 1, wherein in step (A) each of said first DNA molecule, said second DNA molecule and said third DNA molecule is present on a vector, which may be the same or different, comprising a CEN sequence.

10. The method of claim 1, wherein in step (A), at least one of said first DNA molecule, said second DNA molecule and said third DNA molecule is present on a vector, which may be the same or different, and comprises a CEN sequence, and at least one of said first DNA molecule, said second DNA molecule and said third DNA molecule is present on a high copy number vector, which may be the same or different.

11. The method of claim 9, wherein said vector comprising a CEN sequence is a YAC vector.

12. The method of claim 10, wherein said vector comprising a CEN sequence is a YAC vector.

13. The method of claim 10, wherein said high copy number vectoris a YEp plasmid.

14. The method of claim11, wherein said first DNA molecule, said second DNA molecule and said third DNA molecule are present on the same YAC vector.

15. The method of claim 1, wherein said yeast host cell is a member of a genus selected from the group consisting of *Kluveromyces, Saccharomyces, Schizosaccharomyces, Yarrowia* and *Pichia*.

16. A yeast host cell capable of producing a hydroxylated triple helical protein upon culturing, said yeast host cell including:

(i) a first DNA sequence encoding prolyl 4-hydroxylase I-subunit (P4HI) operably linked to a promoter functional in said yeast host cell,

(ii) a second DNA sequence encoding prolyl 4-hydroxylase 9-subunit (P4H9) operably linked to a promoter functional in said yeast host cell, and

(iii) a third DNA sequence encoding a polypeptide or peptide operably linked to a promoter functional in said yeast host cell, wherein said polypeptide or peptide is one which, when hydroxylated, forms said hydroxylated triple helical protein, and wherein said polypeptide or peptide is a synthetic polypeptide or peptide represented by the following formula:

$$(A)_l - (B)_m - [Z] - (C)_o - (D)_p,$$

wherein;

Z is a domain comprising two or more repeat units of the formula:

$$[(E)_q - (GlyXY)_i - (F)_r],$$

wherein;

E and F represent sequences of one or more amino acids, which sequences may vary from repeat unit to repeat unit, and for each repeat unit q and r are each independently selected from 0 and 1, and

i is $\geq 1$ such that domain Z comprises 2 to 1500 GlyXY triplets,

Gly represents glycine, and

X and Y, which may be the same or different, represent an amino acid, and wherein the identity of each amino acid represented by X and Y may vary from GlyXY triplet to GlyXY triplet, but wherein at least one Y of the $(GlyXY)_i$ sequence must be proline,

A and D, which may be the same or different, each represent a polypeptide or peptide domain which optionally comprises a triple helical forming repeating sequence $(GlyXY)_n$, and l and p are each independently selected from 0 and 1,

B and C, which may be the same or different, each represent a polypeptide or peptide domain which is heterologous to collagen proteins and which does not comprise a triple helical forming repeating sequence $(GlyXY)_n$, and m and o are each independently selected from 0 and 1; and

wherein each of said first DNA sequence, said second DNA sequence and said third DNA sequence are replicated, stably retained and segregated by the yeast host cell.

17. The yeast host cell of claim 16, wherein domain Z comprises no more than 10 to 300 GlyXY triplets.

18. The yeast host cell of claim 16, wherein in domain Z, $(GlyXY)_i$ has an amino acid length which is at least three times greater than the combined amino acid length of E and F.

19. The yeast host cell of claim 16, wherein expression of the said P4HI subunit and said and P4HS subunit is controlled in a coordinated manner by a bidirectional promoter.

20. The yeast host cell of claim 19, wherein said bidirectional promoter is yeast GAL1-10 promoter sequence.

21. The yeast host cell of claim 16, wherein said P4HI subunit is an avian P4HI subunit or a mammalian P4HI subunit, and said P4HS subunit is an avian P4HS subunit or a mammalian P4HS subunit.

22. The yeast host cell of claim 21, wherein said mammalian P4HI subunit is human P4HI subunit, and said mammalian P4HS subunit is human P4HS subunit.

23. The yeast host cell of claim 16, wherein each of said second DNA sequence and said third DNA sequence further comprise a DNA sequence encoding a secretion signal such that said P4HS and said polypeptide(s) or peptide(s) are expressed and secreted by said yeast host cell.

24. The yeast host cell of claim 16, wherein each of said first DNA sequence, said second DNA sequence and said third DNA sequence is present on a vector, which may be the same or different, comprising a CEN sequence.

25. The yeast host cell of claim 16, wherein at least one of said first DNA sequence, said second DNA sequence and said third DNA sequence is present on a vector, which may be the same or different, and comprises a CEN sequence, and at least one of said first DNA sequence, said second DNA sequence and said third DNA sequence is present on a high copy number vector, which may be the same or different.

26. The yeast host cell of claim 24, wherein said vector comprising a CEN sequence is a YAC vector.

27. The yeast host cell of claim 25, wherein said vector comprising a CEN sequence is a YAC vector.

28. The yeast host cell of claim 25, wherein said high copy number vector is a YEp plasmid.

29. The yeast host cell of claim 26, wherein said first DNA sequence, said second DNA sequence and said third DNA sequence are present on the same YAC vector.

30. The yeast host cell of claim 16, wherein said yeast host cell is a member of a genus selected from the group consisting of *Kluveromyces, Saccharomyces, Schizosaccharomyces, Yarrowia* and *Pichia*.

31. An hydroxylated triple helical protein comprising a polypeptide or peptide which is a synthetic polypeptide or peptide represented by the following formula:

$$(A)_l - (B)_m - [Z] - (C)_o - (D)_p,$$

wherein;

Z is a domain comprising two or more repeat units of the formula:

$$[(E)_q - (GlyXY)_i - (F)_r],$$

wherein;

E and F represent sequences of one or more amino acids, which sequences may vary from repeat unit to repeat unit, and for each repeat unit q and r are each independently selected from 0 and 1, and

i is $\geq 1$ such that domain Z comprises 2 to 1500 GlyXY triplets,

Gly represents glycine, and

X and Y, which may be the same or different, represent an amino acid, and wherein the identity of each amino acid represented by X and Y may vary from GlyXY triplet to GlyXY triplet, but wherein at least one Y of the $(GlyXY)_i$ sequence must be proline,

A and D, which may be the same or different, each represent a polypeptide or peptide domain which optionally comprises a triple helical forming repeating sequence $(GlyXY)_n$, and l and p are each independently selected from 0 and 1,

B and C, which may be the same or different, each represent a polypeptide or peptide domain which is heterologous to collagen proteins and which does not comprise a triple helical forming repeating sequence $(GlyXY)_n$, and m and o are each independently selected from 0 and 1.

32. The hydroxylated triple helical protein of claim 31, wherein domain Z comprises no more than 10 to 300 GlyXY triplets.

33. The hydroxylated triple helical protein of claim 31, wherein in domain Z, $(GlyXY)_i$ has an amino acid length which is at least three times greater than the combined amino acid length of E and F.

34. A biomaterial or therapeutic product comprising said hydroxylated triple helical protein of claim 32.

35. A biomaterial or therapeutic product comprising said hydroxylated triple helical protein of claim 33.

## ABSTRACT

The invention relates to a method for producing hydroxylated triple helical proteins in yeast host cells by introducing to a suitable yeast host cell, DNA sequences encoding the triple helical protein as well as prolyl 4-hydroxylase (PH4), in a manner wherein the introduced DNA sequences are replicated, stably retained and segregated by the yeast host cells.